

How to preserve digital publications

NEDLIB List of

Terms

Genevieve Clavel-Merrin

Schweizerische Landesbibliothek

The NEDLIB List of Terms

The NEDLIB List of

Terms

Genevieve Clavel-Merrin, Schweizerische Landesbibliothek
Bibliothèque Nationale Suisse

Production Management Bureau D'Arts, Amsterdam

Design Herlof Schürmann, Amsterdam

Printed by Mart.Spruijt, Amsterdam

NEDLIB Report Series Editor Titia van der Werf

Available from Koninklijke Bibliotheek

P.O. Box 90407

2509 LK The Hague

e-mail info@kb.nl

Title The NEDLIB list of terms

ISBN 906259151-5

Author Genevieve Clavel-Merrin/Schweizerische Landesbibliothek – Bibliothèque Nationale Suisse

Date December 2000

Copyright NEDLIB consortium

This study is jointly funded by the European Commission's Telematics for Libraries Programme.

This is report 7 of the NEDLIB Report Series.

content

Introduction	I
List of Terms	3
Contributions	17

Introduction

This glossary defines concepts relevant to NEDLIB. This glossary does not contain names of programmes, file extensions, institutions, explanations of acronyms such as CD-ROM, nor lists of standards.

We found no authoritative source distinguishing clearly between electronic publication and digital publication / document. Sometimes it may be understood from a context that 'electronic publishing' covers digital documents plus video tapes, photos etc. In NEDLIB the terms are considered to be interchangeable, but the term 'electronic publication' is preferred to digital document or electronic document.

This information does not intend to be a generic and exhaustive source, but only to reflect the NEDLIB requirements and boundaries and to clarify the use within NEDLIB of terms which may be considered ambiguous. Many terms taken from the OASIS concept are included for ease of reference. The full OASIS glossary may be found at: <http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html>.

The glossary presents the terms and related descriptions in alphabetical order. When useful, references (including on-line references) of sources related with those descriptions are also provided. The links to online references were valid as of September 19th 2000.

Where possible, only nouns are used as entries in the list and if required related verbs are clarified in the same entry.

List of terms

Access control

The process of determining whether an identity is permitted to perform some action, such as accessing a resource. It requires prior identification and authentication of the identity. A single user may have several identities with differing permissions.

Accession

The process of formally accepting permanent custody of archival records by an archive. Also occasionally used to denote material received.

<http://www.lva.lib.va.us/state/records/manuals/00m%2Dglo.htm>

Acquisitions

The process of selecting and procuring books, periodicals, and other materials by purchase, exchange, gift or deposit; of processing invoices for payment where required; and of keeping the necessary records of these acquired items.

Archival Information Package

An information packaging concept that requires the presence of Content Information and all the associated Preservation Description Information that is needed to preserve the Content Information over the long term. It has associated Packaging Information.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Archival records

Permanent records determined to be of sufficient additional historical, continuing or enduring value as to warrant the added cost of storage, conservation and preservation in an archival facility.

<http://www.lva.lib.va.us/state/records/manuals/00m%2Dglo.htm>

Archive

An organisation that intends to preserve information for access and use by one or more designated communities

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Authentication

A process to establish the authenticity of digital material stored in a Deposit System for Electronic Publications for long-term preservation and access purposes. The authentication process is part of preservation handling processes that alter characteristics of the original publication. Preservation handling processes may alter characteristics of the original document. The authentication process incorporated in the preservation handling process enables such changes to be monitored and authenticity degradation to be measured or established.

Authenticity

Measure or statement establishing how much and in what way an electronic publication stored for long-term preservation and access purposes deviates from the digital original.

Authority control

The process of verifying and authorising the choice of unique access points, such as personal names and subjects, and ensuring that the access points are consistently applied.

<http://info.wgbh.org/upf/glossary.html>

CD-ROM publication

An electronic publication originally distributed on CD-ROM.
Not to be confused with offline batch deposits.

CIP

Cataloguing in Publication

Consumer

The role played by those persons, or client systems, who interact with OAIS services to find preserved information of interest and to access that information in detail.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Content Information

That set of information that is the primary target for preservation.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Context Information

Information that documents the relationships of the Content Information to its environment. This includes why Content Information was created and how it relates to other Content Information objects existing elsewhere.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Data

The representation forms of information. Examples of data include a sequence of bits, a table of numbers, the characters on a page.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Database

A set of data structured to facilitate retrieval and further processing.

Guidelines on best practices for using electronic information / [DLM-Forum on Electronic Records]. 1997
<http://www.ispo.cec.be/dlm/documents/guidelines.html>

Delivery&Capture

Pre-processing interface responsible for receiving deposit publications from the publishers and for submitting them to the DSEP system as a standard submission package [NEDLIB REPORT 6].

Deposit transfer area

Temporary storage area for clearing incoming deposit publications [NEDLIB REPORT 6].

Descriptive Information

That set of information, consisting primarily of Package Descriptions, which is provided to Data Management to support the finding of preserved information by Consumers.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Designated Community

An identification of a set of potential Consumers who should be able to understand a particular set of information.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Digital archive

1. A system ensuring the preservation for future access of electronic documentary resources.

Information paper Digital Services project / National Library of Australia, 16 December 1998.

Appendix-G

<http://www.nla.gov.au/download/dsp/appendices.pdf>

2. Long-term, off-line electronic storage, generally on some form of disc or magnetic tape.

<http://info.wgbh.org/upf/glossary.html>

Digital Library System (DLS)

The ICT-infrastructure and automated library systems that support library business processes, in particular the handling of digital library collections.

Digital object

An object composed of a set of bit sequences.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Digital original

The electronic publication as deposited by a publisher. The original publication may include all of the following: data files, installation and help files, software requirements to read the publication, system requirements to run the software and in some case software and/or hardware necessary to run the publication.

Dissemination Information Package (DIP)

An information package that contains parts or all of one or more Archival Information Packages and that is distributed to the Consumer as requested.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

DSEP

Deposit System for Electronic Publications

Electronic document

See [electronic publication](#)

Electronic publication

A document disseminated in machine-readable form. It includes off-line (physical format) publications distributed on CD-ROM or any other portable medium, and on-line publications stored using digital technology. Some electronic publications are 'born digital' (i.e. are created in digital form) and some are created originally in another form e.g. print on paper, photograph, LP and have subsequently been digitised.

See also [off-line publication](#), [on-line publication](#), [hybrid electronic publication](#)

Electronic publication definition file

The electronic publication definition file specifies the component parts (e.g. metadata files, data files, software files, etc.) and the logical units (table of contents, part/whole relationships) of the publication. This file permits DSEP to perform logical integrity checks [NEDLIB REPORT 6].

Emulation

The process of setting up a system to perform in the same way as another system of a different type in order to run its programs.

Information paper Digital Services project / National Library of Australia, 16 December 1998.

Appendix G

<http://www.nla.gov.au/download/dsp/appendices.pdf>

EP

Electronic publication

Finding aid

Catalogues, classification schemes, descriptive lists, and indexes of various kinds which have been devised to enable information and materials to be located and retrieved.

Fixity Information

This information documents the authentication mechanisms and provides authentication to ensure that the Content Information object has not been altered in an undocumented manner.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Format

The arrangement of data for computer input or output, such as the number and size of data fields in a logical record or the spacing and letter size used in a document.

<http://info.wgbh.org/upf/glossary.html>

Genre

A class or category of documents characterised by properties, belonging to multiple technical and contextual dimensions, and imposing a common set of requirements that impacts on one or more core areas of their management and handling in deposit collections (namely, acquisition, verification, registration, access and preservation).

Handle

A name used to identify items and other Internet resources.

'Identification' : Work Package 2 of Telematics for Libraries project BIBLINK (LB 4034). Glossary <http://hosted.ukoln.ac.uk/biblink/wp2/d2.1/docoo03.htm>

Handle system

The Handle System is a distributed computer system which stores these names and provides the information that is needed to locate and access these items.

'Identification' : Work Package 2 of Telematics for Libraries project BIBLINK (LB 4034). Glossary <http://hosted.ukoln.ac.uk/biblink/wp2/d2.1/docoo03.htm>

Harvesting

The automatic search of www sites, with a view to the retrieval of references to www pages, links and other information.

Hybrid electronic publication

An off-line publication that contains links to on-line material.

Identifier

A character or sequence of characters designed to uniquely designate a document. Examples of identifiers are URN (Uniform Resource Name), DOI (Digital Object Identifier), ISBN (International Standard Book Number), SICI (Serial Item and Contribution Identifier).

Indexing

The creation of a table or list of the contents of a storage medium, file, document, or database, together with keys or references for locating the contents. An index includes physical location information (i.e., which files on which disk) and document identification information (e.g., date, creator, contents).

In full-text indexing, every word from a document is extracted by program and used as an access point to that document. A list of frequently occurring words (a, the it etc.) may be used to exclude certain entries (stop-list). Full-text indexing is usually carried out without human intervention. It may also be termed automatic or key-word indexing.

In subject indexing, a document is examined and its main topics are determined. Words or phrases from a closed, pre-defined list (controlled vocabulary, Subject Heading Language, thesaurus) which are considered to reflect the topic(s) discussed in that document are assigned manually according to rules for their application and number.

Information Object

A Data Object together with optional Representation Information.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Information Package

An information packaging concept that distinguishes Content Information from associated Preservation Description Information where the Preservation Description Information applies to the Content Information and is needed to aid in the preservation of the Content Information. It has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Ingest

An entity containing the services and functions that accepts Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting descriptive Information become established within the OAIS.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Integrity

NEDLIB distinguishes between three levels of integrity:

- physical integrity of digital data. Integrity checking is done to ensure that bits have not been altered and files not been corrupted, during medium refreshing or copying for example.
- logical integrity of an electronic publication, consisting of a set of files. Logical integrity checking is done to ensure all files and the structural relationships between them are not impaired.
- intellectual integrity of an electronic publication. When the content of an electronic publication has been changed through conversion for example, the content is no longer authentic and affects the intellectual integrity of the publication.

Logical publication unit

Electronic publications are defined more in terms of logical and structural boundaries than by physical carrier boundaries. A logical publication unit defines a given granularity level of the electronic publications and is highly dependent on the type of publication (e-journal, web-publication, database-like publication, etc.) [NEDLIB REPORT 6].

This concept is similar to the OAIS concept of Archival Information Unit (AIU).

National Bibliographic Number (NBN)

Unique identifier associated with a deposited publication and its bibliographic description in the National Bibliography. The NBN is issued by a National Bibliographic Agency.

Management

Those who set overall OAIS policy as one component in a broader policy domain.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Metadata

1. Structured information, perhaps contained in an attached header, that describes other resources. Catalogue records for library materials are a common example of metadata. While the resources are interesting to the end user, the metadata is interesting to the people or programs that have to manage the information. Metadata assists in the process of retrieving information by enabling users to initially discover the existence of the information, to locate it and then to determine if it is the information that the user wants. Usually the metadata describes the contents, physical description, location, type and form of the information, and information necessary for management including migration history, expiry dates, security, authentication, and file formats.

<http://info.wgbh.org/upf/glossary.html>

2. Metadata are structured, encoded data that describe characteristics of information-bearing entities to aid in the identification, discovery, assessment, and management of the described entities.

<http://www.ala.org/alcts/organization/ccs/ccda/tf-meta3.html>

Migration

Set of organised tasks designed to achieve the periodic transfer of digital materials from one hardware or software configuration to another, or from one generation of computer technology to a subsequent generation.

The periodic transfer of digital materials from one hardware/software configuration to another or from one generation of computer technology to a subsequent generation. The purpose of migration is to preserve the integrity of digital objects and to retain the ability for clients to retrieve, display, and otherwise use them in the face of constantly changing technology. Migration includes refreshing as a means of digital preservation but differs from it in the sense that it is not always possible to make an exact digital copy or replica of a data base or other information object as hardware and software change and still maintain the compatibility of the object with the new generation of technology.

<http://info.wgbh.org/upf/glossary.html>

see also [Transformation](#)

OAIS-RM

Open Archival Information System Reference Model

Open Archival Information System (OAIS)

An archive, consisting of an organisation of people and systems, that has accepted responsibility to preserve information and make it available for one or more Designated Communities.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Off-line batch deposits

Batch-like deposit of online publications on a physical carrier such as CD-ROM.

Off-line publication

An electronic publication, which is disseminated to its target audience using a digital storage medium (such as a CD-ROM), suitable to be read using a computer or a similar device. It may subsequently be transmitted on-line over a network.

On-line publication

An electronic publication, which is disseminated to its target audience over a computer network such as the Internet.

Packaging&Delivery

Post-processing interface responsible for delivering deposit publications upon request to the library access systems or external systems. It takes care of the request negotiations with DSEP and of all processes needed to unpack a DIP and to make it fit for use by the requesting system [NEDLIB REPORT 6].

Packaging Information

The information used to bind and identify the components of an Information Package.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Preservation

1. The provision of adequate facilities to protect, care for, or maintain records.

<http://info.wgbh.org/upf/glossary.html>

2. Specific measures, individual and collective, undertaken to maintain, repair, restore, or protect records.

<http://info.wgbh.org/upf/glossary.html>

3. Measures aimed at averting damage or reducing the rate of natural, physical or logical ageing of documents (prospective preservation) or the restoration of damaged documents (retrospective preservation).
Preservation may be physical (medium refreshing, medium migration); logical (format conversion, emulation) or intellectual (content preservation).

Preservation Description Information (PDI)

Information necessary to adequately preserve the Content Information and which can be categorised as Provenance, Reference, Fixity, and Context Information

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Preservation Layer Model (PLM)

Model of the different layers of Representation Information, based on the layered Information Model from Annex E of the OAIS-RM and adapted by NEDLIB [NEDLIB REPORT 2]. The PLM can be used as a preservation tool to keep track of all software and hardware dependencies of electronic publications [NEDLIB REPORT 6].

Producer

The role played by those persons, or clients systems, who provide the information to be preserved.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Provenance Information

Information that documents the source of the Content Information (e.g. Publisher) and its history (e.g. version).

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Public records

Any record created or accumulated during the course of public business, including paper, microfilm, electronic, photographic, audio, motion picture or video recordings

Record

Archives:

A specific piece of recorded information generated, collected or received in the initiation, conduct or completion of an institutional or individual activity and that comprises sufficient content, context and structure to provide proof of evidence of that activity.

Guide for Managing Electronic Records from an Archival Perspective / the Committee on electronic records from the International Council on Archives (Feb. 1997)
http://www.ica.org/archives/cer/english_970821.html#top

Database:

A collection of related items of data that is treated as a unit in a database.

<http://www.mbakercorp.com/environmental/whatdata.htm>

Reference Information

Information that identifies, and if necessary describes, one or more mechanisms used to provide assigned identifiers for the Content Information e.g. ISBN, DOI

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Reference platform

Standard computer system configuration designated by the deposit library for installing and running electronic publications. The Reference Platform is suited for most publications appearing on the consumer market, during a given period of time [NEDLIB REPORT 6].

Refreshing

The periodic re-recording of data onto the same support (such as a magnetic tape) or the copy to a new support. Some times also referred as restoration, or transcription.

To copy digital information from one long-term storage medium to another of the same type, with no change whatsoever in the bit-stream (e.g. from a decaying 800bpi tape to a new 800bpi tape, or from an older 5 1/4 inch floppy to a new 5 1/4 inch floppy).

<http://info.wgbh.org/upf/glossary.html>

Registration

The act of recording the acquisition and arrival of a document (electronic or otherwise).

Not to be confused with cataloguing or the preparation and storage of metadata.

Repository

A network-accessible storage system in which digital objects may be stored for possible subsequent access or retrieval.

Representation Information

Information that maps data into more meaningful concepts. An example is the ASCII definition which describes how bits (i.e. Data) are mapped into numbers. Another example is a description of the numbers (i.e. Data) of a table as being the co-ordinates of a location on the Earth measured in East longitude and latitude.

Reference model for an open archival information system (OAIS). April 1999

http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Representation Information Repositories

Repositories containing a common pool of Representation Information, shared by a variety of deposit libraries and other memory organisations. These Repositories provide open standard format specifications, common viewing software, conversion software, interpreters, emulators, etc.

[NEDLIB REPORT 6].

Scenario

The description of a case in a deposit library comprising the identification of a (real or theoretical) related set of digital document genres, the description of their requirements, plus an approach to their management and handling.

Submission Information Package (SIP)

An agreement reached between an OAIS and the Producer that specifies a data model for Data Submission Session. This data model identifies format / contents and the logical constructs used by the Producer and how they are represented on each media delivery or in the telecommunication session.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Surrogate

1. A representation of a document. It may include metadata such as title information, author, date, subject terms, keywords or consist of an abstract. It may be linked to the full-text of the document itself. A surrogate may be generated by a search engine from a document or created manually (i.e. cataloguing).
2. A version that stands for an original – not necessarily an exact reproduction. A copy on microfilm is an example of a surrogate. This use of the word includes lesser versions (e.g., thumbnail images) and reformatted versions (e.g., a digital audio version of an analog recording), but not a part of the whole (e.g., a detail from a photograph).

Guidelines for extending the use of Dublin Core elements to create a generic application integrating all kinds of information resources; Draft 10/1/97 RLG
<http://www.ohio.rlg.org/metawg.html>

For uses of the term in the different senses see:
<http://www.dascorp.com/products/embark/embarkspec>
<http://www.bodley.ox.ac.uk/icsu/mischoppr.htm>
<http://www.ualberta.ca/~nfriesen/536/wwwsrcht.htm>

Transcription

See [Refreshing](#).

Transformation

A Digital Migration in which there is an alternation to the Content Information or PDI of an Archival Information Package. For example, changing ASCII codes to UNICODE in a text document being preserved is a Transformation.

Reference model for an open archival information system (OAIS). April 1999
http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html

Verification

The act of checking the physical and logical consistency of a document (are all bits, files correct; are all files present). Checking of the physical medium aims at testing the integrity of the support and the data stored on it. Content checking is concerned with the logical integrity of a document and the authentication of its source.

Contributions

The content of this document is based on material collected by the authors and on material provided by the following contributors:

BNCF	Claudio Di Benedetto, Giovanni Bergamin
BNF	Sonia Zillhardt
BNP	Fernanda Campos, Almerinda Meireles, Fernando Cardoso, José Luis Borbinha
CSC	Markus Fisher, Peter Schmich
DDB	Hans Liegmann, Joerg Berkemeyer
HUL	Juha Hakala, Inkeri Salonharju
KB	Titia van der Werf, Lex Sijtsma
NBR	Hilde Høgås, Torstein Olsen, Fred Moerman
NSA	Charles Noordam, Eric Burger
SNL	Genevieve Clavel-Merrin, Michel Moret
Mark Bide & Associates	Mark Bide